Task Force For Selecting New Children's Instruments

Synopsis of November 18, 1999 Meeting

A meeting of the Task Force for Selecting New Children's Performance Outcome Instruments was held on Thursday, November 18, 1999, at the Sacramento Airport Host Hotel. A copy of the sign-in sheet of those in attendance is attached to this meeting synopsis. The topics of discussion and the actions that were recommended are highlighted below.

- **Welcoming remarks and Introductions** – Attendees introduced themselves and the agenda was reviewed to identify any additions or changes that should be made. Dave Neilson, DMH, requested that his discussion of "Strategies for Using The Data Across Agencies" could be moved earlier in the lineup, as he needed to leave early.

- **Strategies For Using The Data Across Agencies** – Dave Neilson, helped spark a spirited discussion of how and what data are useful across agencies. It was noted by county staff in attendance that raw data from the Children's Performance Outcome System is not shared with individuals from our partner agencies. Rather, in treatment team meetings the ultimate information that is generated from the instruments that can shed light on the child and his or her functioning as well as on the functioning of the family unit is shared.

  It appears that, at this point, our sister agencies are, in general, not clamoring for our raw, client-level, data in the same way that we would like information from their systems. Fred Hawley, Kern County, stated that this is because they do not have the same operating philosophies as we do. It was stated that the information our sister agencies tend to find more useful comes in the form of annual reports that can shed light on children and families and the services they receive.

  Jim Higgins, DMH, suggested that this was an important insight for him because he had been thinking that our goal was a system that had raw data on every client from every service system that could be accessed for research and evaluation purposes. While some counties may in fact be able to do this, the State as a whole is apparently very far from having this be the rule.

  Also, in terms of sharing data across agencies, it was noted that there are technical problems involved in sharing data within our own agencies and that the challenge of implementing a vision of shared data across agencies and software platforms remains problematic. County staff reported that in some counties, some of our sister agencies simply don't have the data to share (e.g., educational data).

  The conclusion is that we must not lose sight of our goal of collecting the smallest amount of data that will answer our questions, help us improve our systems and

services, and which can facilitate improvement in the entire system of care through inter-agency collaboration.

- **Review of Data Collected From The Statewide Survey On The Children's Performance Outcome System** – Sherrie Sala-Moore, DMH, conducted a statewide survey to collect feedback from all participants in the Children's Performance Outcome System regarding how well they felt it was working and, if it should be changed, how. Following are excerpts from the survey analysis:

  - Out of 695 valid responses, 2.7% claimed to be mental health directors, 11.8% were children's coordinators or program administrators, 6.5% were quality managers, 1.9% worked in Information Technology departments, 66.2% were children's clinicians, 1.2% were parents or representatives of consumer groups, and 9.8% were "other".
  - When asked whether all, parts, or none of the system should be changed, 43.8% wanted the entire system replaced, 39.5% felt that at least part of the system should be replaced, and 16.7% recommended no changes.
  - Children's clinicians, children's evaluators, and children's program coordinators were the groups who most wanted the entire system changed.
  - Mental health directors and information technology staff were the people who most wanted the system to stay the same with no changes. It is important to note, however, that this assertion is based on the relatively small number of respondents to the survey from these groups.
  - In terms of exactly what changes should be made to the system, the following recommendations were made:

    - ✓ Change everything (44%)
    - ✓ Change only the CBCL and YSR (12%)
    - ✓ Change only the CAFAS (7%)
    - ✓ Change only the CSQ-8 (9%)
    - ✓ Change the CAFAS, CBCL and YSR but keep the CSQ-8 (5%)
    - ✓ Change the CBCL, YSR and CSQ-8 but keep the CAFAS (6%)
    - ✓ Don't change anything, the system works well (5%)
    - ✓ Don't change anything, it would be too difficult (8%)
    - ✓ Don't change anything, no reason given (4%)

  - How did people feel about aspects of the system other than the instruments themselves?

    - ✓ How much time it takes to complete the forms – 33.7% were very dissatisfied, 32% were somewhat dissatisfied, 11.6% were neutral, 15.5% were somewhat satisfied, 4.9% were very satisfied, and 2.3% had no opinion.
    - ✓ How easy are the instruments to read and understand – 15.6% were very dissatisfied, 28.4% were somewhat dissatisfied, 17.4% were neutral, 26.9% were somewhat satisfied, 10.4% were very satisfied, and 1.3% had no opinion.

- ✓ How valuable is the data generated from the current instruments for treatment planning – 29.4% were very dissatisfied, 22.3% were somewhat dissatisfied, 14.5% were neutral, 22.3% were somewhat satisfied, 7.5% were very satisfied, and 4% had no opinion.
- ✓ How valuable is the data for quality management – 29.3% were very dissatisfied, 18.8% were somewhat dissatisfied, 25.3% were neutral, 14.3% were somewhat satisfied, 4% were very satisfied, and 8.3% had no opinion.
- ✓ How useful are the reports and profiles that result from the instruments – 30.3% were very dissatisfied, 25.5% were somewhat dissatisfied, 15.9% were neutral, 13.7% were somewhat satisfied, 5.5% were very satisfied, and 9% had no opinion.
- ✓ How easy are the instruments and their data integrated with Information Management Systems – 30.7% were very dissatisfied, 17.2% were somewhat dissatisfied, 20.3% were neutral, 6.6% were somewhat satisfied, 3.1% were very satisfied, and 22.2% had no opinion.
- ✓ In terms of cultural sensitivity or neutrality, how appropriate did people think the instruments were for various cultures – 14.4% were very dissatisfied, 20.4% were somewhat dissatisfied, 30.8% were neutral, 17.5% were somewhat satisfied, 8.7% were very satisfied, and 8.7% had no opinion.
- ✓ How satisfied were people with the extent to which the instruments focused on strengths and not just problems – 18.3% were very dissatisfied, 25.3% were somewhat dissatisfied, 23.9% were neutral, 19.2% were somewhat satisfied, 6.5% were very satisfied, and 6.8% had no opinion.
- ✓ How suitable are the instruments for our target population – 20.6% were very dissatisfied, 22.0% were somewhat dissatisfied, 18.5% were neutral, 25.0% were somewhat satisfied, 9.5% were very satisfied, and 4.4% had no opinion.
- ✓ How satisfied were people with the prospects of continuing to use the current system in the long term – 32.3% were very dissatisfied, 23.4% were somewhat dissatisfied, 20.2% were neutral, 13.5% were somewhat satisfied, 4.0% were very satisfied, and 6.7% had no opinion.

- In terms of priorities, we asked people what they felt were the most important criteria for evaluating the existing system and any potential alternatives. In order to their ratings of importance, the top five criteria that they would use to judge a system were:

  1. The system must include data collected from multiple informants (not just the clinician).
  2. Psychometric validity and reliability of the instruments.
  3. The instruments must be short and easy to administer.
  4. Low cost instruments (public domain preferred).
  5. The data that are collected must be cost effective (value of the data per time and cost).

- Sherrie also conducted an informal survey of counties to get a rough estimate of the costs associated with administering the children's instruments. This is a

question that had been asked many times and by many people, but answers have been unavailable. It is important to note that the survey was not scientific and involved asking county staff to make rough estimates of clerical time and clinician time involved in administering and scoring the instruments on a per client basis. Additionally, staff hourly costs were gathered. The result was an estimate of the low, average and highest costs associated with administering the instruments. Sherrie then used DMH's Client Data System (CDS) data file to estimate the county's children's performance outcome target population count and then computed a total cost for instrument administration. The cost estimates did not include overhead such as information technology infrastructure development and maintenance costs. The costs per client ranged from a low of $10 to a high of $130 to administer the set of instruments with a typical average county cost of $40.

- **Review of The California Mental Health Planning Council's (CMHPC) comparison of the tentatively selected children's pilot instruments and their coverage of the CMHPC domains** – Ann Arneill-Py reviewed an evaluation that she conducted of the Ohio Scales, the Behavioral and Emotional Rating Scale (BERS), Columbia Impairment Scale (CIS), and the Colorado Client Assessment Record. Based on Ann's careful analysis, the best instrument (instrument set) that we found was the Ohio Scales. These scales include a clinician rating, parent rating, and child rating. Additionally, we can add our own risk and categorical variables to the clinician form of the Ohio scales. They are very short, have been validated, and seem to be quite appropriate for the level of analysis that we want to do.

  The Colorado Client Assessment Record (C-CAR) has a lot to offer as well. However, there is a concern that it is too much like a county's standard assessment information and so it would appear redundant. Ann Arneill-Py also had not been able to obtain any specific validation information from the C-CAR's authors in spite of repeated attempts.

  Ann was not particularly impressed with either the BERS or the CIS. The CIS was especially deficient in that it only results in a single global score similar to the Global Assessment of Functioning (GAF).

- **Developing a Draft Protocol for conducting a pilot of the instruments selected for piloting** – what was to be a discussion of the procedures for conducting a pilot test of the Ohio Scales and any other instruments that could be identified that would meet the established evaluation criteria morphed into a discussion of the broader issues of exactly what constitutes an effective and valid performance outcome measurement methodology.

  The current methodology applied to the Children's Performance Outcome System involves administering instruments to the same clients over time in a repeated measures design. This methodology, if it could be made to work properly, would allow us to answer a variety of important questions. Still there would be some

limitations. For example, long term clients (those for whom multiple year data would be available) are likely to show improvement early in their treatment program. However, the trajectory of improvement tends to level off with those who are longest in the system not exhibiting measurable change from year to year.

Additional problems with the longitudinal approach tend toward the fact that the sample of clients for whom we have longitudinal data becomes increasingly biased over time. Following is an illustration of the problem:

1. We only have data at time 1 (either intake or for existing clients those who came up for their first annual review during the first year of implementation) for clients who a) did not refuse to complete the forms or b) clinicians did not fail to try and complete the forms. This means that our sample at time one is biased from the start. This is evidenced by the fact that almost no county has managed to include more than about 60% of their target population in the children's system.

2. At time 2 the number of people who complete the forms drops dramatically. This could be because, a) now that the client is receiving services there is reduced willingness to complete the forms and so they refused, b) the client dropped out of services because he, she, or the family as a whole decompensated to the point where they could not seek services, c) the client got better and so the family decided services were no longer needed and they chose not to officially discharge, d) the client left the county without officially discharging from services, or e) the clinician simply chose not to administer the instruments. Therefore, the clients for whom we have time 1 AND time 2 data is even more restricted and likely much less representative of the target population as a whole. An example of this is that the average drop off rate in the number of people for whom time 2 data are available is around 40%.

3. At time 3 and each subsequent year, the drop off rate in data becomes extreme. One county's Children's Evaluator explained that, out of around 10,000 children from whom time 1 data were available, when looking at five years, the number of children for whom longitudinal data were available was only around 250. Under such cases, given the amount of money and effort that goes into collecting longitudinal data, the amount of usable information that is available under this model is severely restricted.

The discussion continued with the notion that there may be better models for collecting, analyzing and using outcome data. Several possible models are noted below:

1. Rather than longitudinal, repeated measures methodology, utilize a point-in-time cross-sectional methodology. This would involve administering a set of state instruments one or two times a year to every child from the target population who receives services. This would capture new clients as well as ongoing clients who receive a service. The data would then be analyzed and clients classified

according to specific characteristics and risk factors. Each county would evaluate their data in order to identify service areas where improvement is either needed or desired and set goals that they would like to achieve. The county would then implement an intervention designed to improve the targeted service. The following year, when the data is again collected, the county and state would evaluate the data to see if the intervention was successful and improvement occurred. If not, the county may either implement a new intervention or target another area of their services where their improvement efforts might be more effective.

Additionally, during the time between cross-sectional data collection months, more carefully designed and controlled longitudinal studies would be completed on small and more representative samples of children. These children could be followed even if they drop out of services as well as tracked regarding their usage of other systems. Such small studies would be able to answer the questions that a cross-sectional study may be less effective at answering.

Jim Higgins, DMH, even said that, in his mind, the best approach would be to have the State duplicate and prepare the forms, mail them to the county where they would be administered during a given month. After which, the forms would be mailed to DMH for scanning and data cleanup. Once the data were cleaned, they would be exported to each county's DMH bulletin board site or ITWS data directory where they could then be downloaded by the county for local analysis. The result of this would be the freeing of counties from having to purchase special software and hardware to score and manage the data, eliminate the need to track clients across programs just to comply with State outcome reporting requirements, and free evaluation staff to do more work directly tied to using the data for local analysis and quality improvement efforts.

2. Another method that was mentioned would be to administer the forms only to people who are early in treatment since these are the ones who are likely to change. While this would measure change, a down side is that the requirements for performance outcomes seems to center around focusing on those children who are long term users of the system. This has been the hallmark of how DMH, CMHPC, and the California Mental Health Director's Association have defined the target population. However, this method would also have advantages in terms of system complexity.

3. A third recommendation was to focus, rather than on the county as a whole, on specific providers. While this would reduce the scale of the system, county staff noted that it would be incredibly complex to manage.

Ultimately, these ideas each have strengths and weaknesses. Each is designed to answer certain questions. There are some questions a longitudinal approach, when done correctly, can answer that a cross-sectional study can't. Additionally, cross-

sectional studies are less complex, provide more timely data and answer questions that a longitudinal approach can't.

The final result was that the Task Force will dedicate a large portion of the next meeting (scheduled for January 24, 2000) to discussing the pros and cons of various methodologies. Specifically, each methodology will be evaluated in terms of exactly what kinds of questions each is intended to answer as well as how robust it is given the limitations of the kinds of data that are available to us. To the extent possible, some experts in research methodology will be identified and invited to attend to help in this important discussion.

Upon completion of the discussion of methodologies, a research protocol will be designed that is intended to test not only the recommended pilot instruments, but also the methodologies that may be used in order to find the approach that works best for California's public mental health system and the consumers whom it serves.

- **Next Meeting** - Sacramento Airport Host Hotel, American Room
  January 24, 2000
  10:00 AM – 3:00 PM